

KIAA0313

- Clone 7.7 exhibits 75% identity and 95% similarity of human clone KIAA0313.

GGKDVSAEAESSMVPVTTTEAKVPMPAHIAVTPSTTKGLIARKEGRYREPPPTPPGYN
GGKDVSI EAESSLTSVTTEETKVPMPAHIAVASSTTKGLIARKEGRYREPPPTPPGYI
***** *****. *****

GGIPIADFPFGCHPARKPPDYNVALQSRMVARPTEAPAG--QTPP-AAAASRPGSKPQ
GIIPIITDFPEGHSHPARKPPDYNVALQSRMVARSSDTAGPSVQQPHGHTSSRPVNKPQ

WHKPSDADPRLAPFQAGFAGAEDEDEQVSAV (SEQ ID NO: 28)
 WHKPNESDPRLAPYQSQGFSTEDEDEQVSAV (amino acid nos. 1348 to 1499 of SEQ ID NO: 2)



4618 b.p.

1/1 31/11
atg aaa tca cta gca atc cca gct aac cat gga gtt atg ggc cag cag gag aaa cac tca
M K P L A I P A N H G V M G Q Q E K H S
61/21 91/31
ctt cct gca gat ttc aca aaa ctg cat ctt act gac agt ctc cac cca cag gtg acc cac
L P A D F T K L H L T D S L H P Q V T H
121/41 151/51
gtt tct tct agc cat tca gga tgt agt atc act agt gat tct ggg agc agc agt ctt tct
V S S S H S G C S I T S D S G S S S L S
181/61 211/71
gat atc tac cag gcc aca gaa agc gag gct ggt gat atg gac ctg agt ggg ttg cca gaa
D I Y Q A T E S E A G D M D L S G L P E
241/81 271/91
aca gca gtg gat tcc gaa gac gac gac gat gaa gaa gac att gag aga gca tca gat cct
T A V D S E D D D D E E D I E R A S D P
301/101 331/111
ctg atg agc agg gac att gtg aga gac tgc cta gag aag gac cca att gac cgg aca gat
L M S R D I V R D C L E K D P I D R T D
361/121 391/131
gat gac att gaa caa ctc ttg gaa ttt atg cac cag ttg cct ttt gcc aat atg aca
D D I E Q L L E F M H Q L P A F A N M T
421/141 451/151
atg tca gtg agg cga gaa ctc tgt gct gtg atg gtg ttc gca gtg gtg gaa aga gca ggg
M S V R R E L C A V M V F A V V E R A G
481/161 511/171
acc ata gtg tta aat gat ggt gaa gag ctg gac tcc tgg tca gtg att ctc aat gga tct
T I V L N D G E E L D S W S V I L N G S
541/181 571/191
gtg gaa gtg act tat cca gat gga aaa gca gaa ata ctg tgc atg gga aat agt ttt ggt
V E V T Y P D G K A E I L C M G N S F G
601/201 631/211
gtc tct cct acc atg gac aaa gaa tac atg aaa gga gtg atg aga aca aag gtg gat gac
V S P T M D K E Y M K G V M R T K V D D
661/221 691/231
tgc cag ttt gtc tgc ata gcc cag caa gat tac tgc cgt att ctc aat caa gta gaa aag
C O F V C I A O O D Y C R I L N O V E K
721/241 751/251
aac atg caa aaa gtt gaa gag gaa gga gag att gtt atg gtg aaa gaa cac cga gaa ctt
N M O K V E E E G E I V M V K E H R E L
781/261 811/271
gat cga act gga aca aga aag gga cac att gtc atc aag ggt acc tca gaa agg tta aca
D R T G T R K G H I V I K G T S E R L T
841/281 871/291
atg cat ttg gtg gaa gag cat tca gta gta gat cca aca ttc ata gaa gac ttt ctg ttg
M H L V E E H S V V D P T F I E D F L L
901/301 931/311
acc tat agg act ttt ctt tct agc cca atg gaa gtg ggc aaa aag tta ttg gag tgg ttt
T Y R T F L S S P M E V G K K L L E W F
961/321 991/331
aat gac cag agc ctc agg gat aag gtt aca cgg gta gta tta ttg tgg gta aat aat cac
N D P S L R D K V T R V V L L W V N N H
1021/341 1051/351
ttc aat gac ttt gaa gga gat cct gca atg act cga ttt tta gaa gaa ttt gaa aac aat
F N D F E G D P A M T R F L E E F E N N

cNMP-BD

REM



1081/361 1111/371
 ctg gaa aga gag aaa atg ggt gga cac cta agg ctg ttg aat atc gcg tgt gct gct aaa
 L E R E K M G G H L R L L N I A C A A K
 1141/381 1171/391
 gca aaa aga aga ttg atg acg tta aca aaa cca tcc cga gaa gct cct ttg cct ttt atc
 A K R R L M T L T K P S R E A P L P F I
 1201/401 1231/411
 tta ctt gga ggc tct gag aag gga ttt gga atc ttt gtt gac agt gta gat tca ggt agc
 L L G G S E K G F G I F V D S V D S G S
 1261/421 1291/431
 aaa gca act gaa gca ggc ttg aaa cgg ggg gat cag ata tta gaa gta aat ggc caa aac
 K A T E A G L K R G D O I L E V N G Q N
 1321/441 1351/451
 ttt gaa aac att cag ctg tca aaa gct atg gaa att ctt aga aat aac aca cat tta tct
 F E N I O L S K A M E I L R N N T H L S
 1381/461 1411/471
 atc act gtg aaa acc aat tta ttt gta ttt aaa gaa ctt cta aca aga ttg tca gaa gag
 I T V K T N L F V F K E L L T R L S E E
 1441/481 1471/491
 aaa aga aat ggt gcc ccc cac ctt cct aaa att ggt gac att aaa aag gcc agt cgc tac
 K R N G A P H L P K I G D I K K A S R Y
 1501/501 1531/511
 tcc att cca gat ctt gct gta gat gta gaa cag gtg ata gga ctt gaa aaa gtg aac aaa
 S I P D L A V D V E Q V I G L E K V N K
 1561/521 1591/531
 aaa agt aaa gcc aac act gtg gga gga agg aac aag ctg aaa aag ata ctc gac aag act
 K S K A N T V G G R N K L K K I L D K T
 1621/541 1651/551
 cgg atc agt atc ttg cca cag aaa cca tac aat gat att ggg att ggt cag tct caa gat
 R I S I L P Q K P Y N D I G I G Q S Q D
 1681/561 1711/571
 gac agc ata gta gga tta agg cag aca aag cac atc cca act gca ttg cct gtc agt gga
 D S I V G L R Q T K H I P T A L P V S G
 1741/581 1771/591
 acc tta tca tcc agt aat cct gat tta ttg cag tca cat cat cgc att tta gac ttc agt
 T L S S S N P D L L Q S H H R I L D F S
 1801/601 1831/611
 gct act cct gac ttg cca gat caa gtg cta agg gtt ttt aag gct gat cag caa agc cgc
 A T P D L P D O V L R V F K A D O Q Q S R
 1861/621 1891/631
 tac atc atg atc agt aag gac act aca gca aag gaa gtg gtc att cag gct atc agg gag
 Y I M I S K D T T A K E V V I O A I R E
 1921/641 1951/651
 ttt gct gtt act gcc acc ccg gat caa tat tca cta tgt gag gtc tct gtc aca cct gag
 F A V T A T P D O Y S L C E V S V T P E
 1981/661 2011/671
 gga gta atc aaa caa aga aga ctt cca gat cag ctt tcc aaa ctt gca gac aga ata caa
 G V I K O R R L P D O L S K L A D R I O
 2041/681 2071/691
 ctg agt gga agg tat tat ctg aaa aac aac atg gaa aca gaa act ctt tgt tca gat gaa
 L S G R Y Y L K N N M E T E T L C S D E
 2101/701 2131/711
 gat gct cag gag ttg ttg aga gag agt caa att tcc ctc ctt cag ctc agc act gtg gaa
 D A Q E L L R E S Q I S L L O L S T V E
 2161/721 2191/731
 gtt gca aca cag ctc tct atg cga aat ttt gaa ctc ttt cgc aac att gaa cct act gaa
 V A T O L S M R N F E L F R N I E P T E

PDZ

RA

Figure 3C



2221/741 2251/751
tat ata gat gat tta ttt aaa ctc aga tca aaa acc agc tgt gcc aac ctg aag aga ttt
Y I D D L F K L R S K T S C A N L K R F
2281/761 2311/771
gaa gaa gtc att aac cag gaa aca ttt tgg gta gca tct gaa att ctc aga gaa aca aac
E E V I N Q E T F W V A S E I L R E T N
2341/781 2371/791
cag ctg aag agg atg aag atc att aag cat ttc atc aag ata gca ctg cac tgt agg gaa
Q L K R M K I I K H F I K I A L H C R E
2401/801 2431/811
tgc aag aat ttt aac tca atg ttt gca atc atc agt ggc cta aac ctg gca cca gtg gca
C K N F N S M F A I I S G L N L A P V A
2461/821 2491/831
aga ctg cga acg acc tgg gag aaa ctt ccc aat aaa tac gaa aaa cta ttt caa gat ctc
R L R T T W E K L P N K Y E K L F O D L
2521/841 2551/851
caa gac ctg ttt gat cct tcc aga aac atg gca aaa tat cgt aat gtt ctc aat agt caa
Q D L F D P S R N M A K Y R N V L N S Q
2581/861 2611/871
aat cta caa cct ccc ata atc cct cta ttc cca gtt atc aaa aag gat ctc acc ttc ctt
N L O P P I I P L F P V I K K D L T F L
2641/881 2671/891
cac gaa gga aat gac tca aaa gta gac ggg ctg gtc aat ttt gag aag cta agg atg att
H E G N D S K V D G L V N F E K L R M I
2701/901 2731/911
gca aaa gaa att cgt cac gtt ggc cga atg gct tca gtg aac atg gac cct gcc ctc atg
A K E I R H V G R M A S V N M D P A L M
2761/921 2791/931
ttc agg act cgg aag aag aaa tgg cgg agt ttg ggg tct ctc agc cag ggt agt aca aat
F R T R K K K W R S L G S L S Q G S T N
2821/941 2851/951
gca aca gtg cta gat gtt gct cag aca ggt ggt cat aaa aag cgg gta cgt cgt agt tcc
A T V L D V A Q T G G H K K R V R R S S
2881/961 2911/971
ttt ctc aat gcc aaa aag ctt tat gaa gat gcc caa atg gct cga aaa gtg aag cag tac
F L N A K K L Y E D A Q M A R K V K O Y
2941/981 2971/991
ctt tcc aat ttg gag cta gaa atg gac gag gag agt ctt cag aca tta tct ctg cag tgt
L S N L E L E M D E E S L O T L S L Q C
3001/1001 3031/1011
gag cca gca acc aac aca ttg cct aag aat cct ggt gac aaa aag cct gtc aaa tcc gag
E P A T N T L P K N P G D K K P V K S E
3061/1021 3091/1031
acc tct cca gta gct cca agg gca ggg tca caa cag aaa gct cag tcc ctg cca cag ccc
T S P V A P R A G S Q Q K A Q S L P Q P
3121/1041 3151/1051
cag cag cag cca cca cca gca cat aaa atc aac cag gga cta cag gtt ccc gcc gtg tcc
Q Q Q P P P A H K I N Q G L Q V P A V S
3181/1061 3211/1071
ctt tat cct tca cgg aag aaa gtg ccc gta aag gat ctc cca cct ttt ggc ata aac tct
L Y P S R K K V P V K D L P P F G I N S
3241/1081 3271/1091
cca caa gct tta aaa aaa att ctt tct ttg tct gaa gaa gga agt ttg gaa cgt cac aag
P Q A L K K I L S L S E E G S L E R H K
3301/1101 3331/1111
aaa cag gct gaa gat aca ata tca aat gca tct tgg cag ctt tct tct cct cct act tct
K Q A E D T I S N A S S Q L S S P P T S

CDC25

Insertion
Unique to
GRF4

Continuation of the
CDC25 domain

Figure 3D



3361/1121 3391/1131
cca cag agt tct cca agg aaa ggc tat act ttg gct ccc agt ggt act gtg gat aat ttt
P Q S S P R K G Y T L A P S G T V D N F
3421/1141 3451/1151
tca gat tct ggt cac agt gaa att tct tca cga tcc agt att gtt agc aat tct tct ttt
S D S G H S E I S S R S S I V S N S S F
3481/1161 3511/1171
gac tca gtg cca gtc tca ctg cac gat gag agg cgc cag agg cat tct gtc agc atc gtg
D S V P V S L H D E R R Q R H S V S I V
3541/1181 3571/1191
gaa aca aac cta ggg atg ggc agg atg gag agg cgg acc atg att gaa cct gat cag tat
E T N L G M G R M E R R T M I E P D Q Y
3601/1201 3631/1211
agc ttg ggg tcc tat gca cca atg tcc gag ggc cga ggc tta tat gct aca gct aca gta
S L G S Y A P M S E G R G L Y A T A T V
3661/1221 3691/1231
att tct tct cca agc aca gag gaa ctt tcc cag gat cag ggg gat cgc gcg tca ctt gat
I S S P S T E E L S Q D Q G D R A S L D
3721/1241 3751/1251
gct gct gac agt ggc cgt ggg agc tgg acg tca tgc tca agt ggc tcc cat gat aat ata
A A D S G R G S W T S C S S G S H D N I
3781/1261 3811/1271
cag acg atc cag cac cag aga agc tgg gag act ctt cca ttc ggg cat act cac ttt gat
Q T I Q H Q R S W E T L P F G H T H F D
3841/1281 3871/1291
tat tca ggg gat cct gca ggt tta tgg gca tca agc agc cat atg gac caa att atg ttt
Y S G D P A G L W A S S S H M D Q I M F
3901/1301 3931/1311
tct gat cat agc aca aag tat aac agg caa aat caa agt aga gag agc ctt gaa caa gcc
S D H S T K Y N R Q N Q S R E S L E Q A
3961/1321 3991/1331
cag tcc cga gca agc tgg gcg tct tcc aca ggt tac tgg gga gaa gac tca gaa ggt gac
Q S R A S W A S S T G Y W G E D S E G D
4021/1341 4051/1351
aca ggc aca ata aag cgg agg ggt gga aag gat gtt tcc att gaa gcc gaa agc agt agc
T G T I K R R G G K D V S I E A E S S S
4081/1361 4111/1371
cta acg tct gtg act acg gaa gaa acc aag cct gtc ccc atg cct gcc cac ata gct gtg
L T S V T T E E T K P V P M P A H I A V
4141/1381 4171/1391
gca tca agt act aca aag ggg ctc att gca cga aag gag ggc agg tat cga gag ccc ccg
A S S T T K G L I A R K E G R Y R E P P
4201/1401 4231/1411
ccc acc cct ccc ggc tac att gga att ccc att act gac ttt cca gaa ggg cac tcc cat
P T P P G Y I G I P I T D F P E G H S H
4261/1421 4291/1431
cca gcc agg aaa ccg ccg gac tac aac gtg gcc ctt cag aga tct cgg atg gtc gca cga
P A R K P P D Y N V A L Q R S R M V A R
4321/1441 4351/1451
tcc tcc gac aca gct ggg cct tca tcc gta cag cag cca cat ggg cat ccc acc agc agc
S S D T A G P S S V Q Q P H G H P T S S
4381/1461 4411/1471
agg cct gtg aac aaa cct cag tgg cat aaa ccg aac gag tct gac ccg cgc ctc gcc cct
R P V N K P Q W H K P N E S D P R L A P

PY motifs

Figure 3E



4441/1481

4471/1491

tat cag tcc caa ggg ttt tcc acc gag gag gat gaa gat gaa caa gtt tct gct gtt tga
Y Q S Q G F S T E E D E D E Q V S A V *

4501/1501

4531/1511

PDZ binding motif

ggc aca gac ttt tct gga agc aga gcg agc cac ctg aaa gga gag cac aag aag acg tcc
G T D F S G S R A S H L K G E H K K T S

4561/1521

4591/1531

tga gca tlg gag cct tgg aac tca cat tct gag gac ggt gga cca gtt tgc ctc ctt c
* A L E P W N S H S E D G G P V C L L

(SEQ ID NO: 1)

(SEQ ID NO: 2)

Figure 3F

hGRF4 SQISLLQLSTVEVATQLSMRNFELFRNIEPTYIDDLF---KLRSKTSCANLKRFEVIN
dGRF4 SNVHFLHLNAYELAIQTLTQDFANFROIESTEYVDLF---ELRSRYGVPMLSKFAELVN
hEpac SAEGLDLVSAKDLAQLTQHDWSLFNSIHQVELIHYVLGPOHLRDVT-TANLERFMRRFN
mRasGRF2 KAECFETLSAMELAEQITLLDHI VFERSIPYEELFGQG--MKLDKNERTPYIMKTSQHFN
dsOS DEITLLTLHPLELARQLTLLFEFYHKNVKPSSELVGSPWT--KKDKEVKSPNLLKIMKHHT
hRasGRP VSLFDHLPEELS EHLTYLEFKSFRRISDFSQNYLVN---SCVKENPTMERSIALCN

. . : : : : . : :

hGRP4 QETFWVASEILRETNQLKRMKIIKHFIKIALHCRECKNFNSMFAIISGLNLAPVARLRTT
dGRP4 REMFWVVSIEICAEHNIVRRHKIVQKFIKIAHCKEKNFNSHFAIVSGLGHGAVSRLROT
hEpac ELOYWVATELCLCPVGPRAQLLRKFIKLAHLKEQKNLNSFFAVMFGLSNSAISRLAHT
mRasGRF2 EMSNLVASQIMNYADISSRPNAIEKWAVADICRCLHNYNGVLEITSALNRSPIYRLKKT
dSOS NVTRWIEKSITEAENYEERLAIHQRAIEVHVHLELNNFNGILSIVAAMGTASVYRLRWT
hRasGRP GISQWVQLMVLSRPTPOLRAEVFIKFIQVAQKLHQLONFNTLMAVIGGLCHSSISRLKET

```
hGRF4      WEKLPNKYEKLFQDLQDLDPDSRNMAKYRNVLSNQL--QPPIIPLFPPVIKKDLTFLHEGN
dGRF4      WEKLPSKYQRLFNDLQDLMDPSRNSKYRQLVSAELLAQHPIIPFPYIVKKDLTFIHLGN
hEpac      WERLPHKVRKLYSALERLLDPSWNERVYR--LALAKLS--PPVIPFMPLLLKDMTFIHEGN
mRasGRF2   WAKVSKQTKALMDKLOXTVSSEGFKNL--ETLKNCN--PPAVPYLGMNLTDLAFIEGT
dSOS       FQGLPERYRKFLFEECRELSDHLLKKYQER--L--RSIN--PPCVPFPGRYLTNIIHLEEGN
hRasGRP     SSHVPHEINKVLGEMTELLSSSRNYDNYR--RAYGECT--DFKIPILGVHLKDLISLYEAH
```

```

hGRF4      DSKVDG--LVNFEKLRMIAKEIRHVGRMASVNMDPALMFRTRKKK----WRLSGLSLQGS
dGRF4      DTRVDG--LINF EKLRMLAKEVRL LTHMCSSPYDLLSILELGQSPSNALFSLNQMSASQ
hEpac      HTLVEN--LINF EKMRMMARAARMLHBCRSHNPVPLSPLRSRVSHLHEDSQVARI STCSE
mRasGRF2   PNFT EEG-LVNFSKMRMISHIIIREIROFOQTAYRIDQP-----
dsOS        PDLLANTELINFSKRRKVAEII GEIQYQONQPYCLNEES-----
hRasGRP     PDYLEDG-KVN VHKLLALYNHISELVQLQEVAPPLEANKDL-----
            : * : : :

```

```

hGRF4      TNATVLVDVAQTGG----HKKRVRRSSFLNAKKLYEDAQMARKVKOYLSNLEEMDEESLQ
dGRF4      SNAAAGTVIAANAGQATIKRRKKSTAAPNPKMFEEAQMVRRVKAYLSNLSKILSDEDLLH
hEpac      -----QSLSTRSPASTWAYVQOLKVIDNQRELS
mRasGRF2   -----KVIOYLLDKALVIDEDSLY
dSOS       -----TIROFFEQLDPFNGLSDFKQMSDYLY
hRasGRP     -----VHLLTSLDLYYTEDEIY

```

Figure 4

hGRF4	TLSLQCEPATNT	(SEQ ID NO: 2)
dGRF4	RFSLECEPAHGS	(SEQ ID NO: 7)
hEpac	RLSRELEP----	(SEQ ID NO:8)
mRasGRF2	ELSLKIEPRLPA	(SEQ ID NO: 9)
dSOS	NESLRIEPRGCK	(SEQ ID NO: 10)
hRasGRP	ELSYAREPNHR	(SEQ ID NO: 11)



Figure 5: GRF4-REM domain

CDC25	-IRGGTKEALIEHLT-SHELVDAAFNVTMLITFRSILT-TREFFYALTYRY-	(SEQ ID NO: 12)
Sos_mouse_	-IKGGTVVKLIERLT-YHMYADPNF-VRTEFLTYRSFCK-PQELNLLIERFE	(SEQ ID NO: 13)
RasGEF_aimless_	VVKFASLNKLVEHLT-HDSKHDLQFLKTFMLTYQSFT-PEKLMSKLQORY-	(SEQ ID NO: 15)
GRF2_mouse_	-IRYASVEALLERLT-DLRFLSIDFLNTFLHTYRIFTT-ATVVLAKLSDIY-	(SEQ ID NO: 14)
GRF4	-IKG-TSERLTMHLVEEHSVVDPTFIEDFLTYRTFLSSPMEVGKKLLEWFN	(SEQ ID NO: 2)
	:: : * : * * :	



Figure 7

PDZ domain

hGRF4	LTKPSREAPLPFILLGGSEK-----GFGIFVDSVDSGSKATEAG-LKRGDQIL	
dGRF4	LTRSSRDEPLNFRIVGGYELRGVAIATGNAAVGIYISHVEPGSKAQDVG-LKRGDQIH	
hPTP-BAS-1	NLKKDAKYGLGFQIIGGEKMGRL-----DLGIFISSVAPGGPADLDGCLKPGDRLI	
hPSD-95	IVIHRGSTGLGFNIVGGEDG-----EGIFISFILAGGPADLSGELRKGDQIL	
rLin-7-C	VELPKTEEGLGFNIMGKEQ-----NSPIYISRIIPGGIADRHGGLKRGDQLL	
hDGL	VKVQKGSEPLGISIVSGEKG-----GIYVSKVTVGSIAHQAG-LEYGDQLL	
	. * : : : * .	* : : : * . * . * . * .
hGRF4	EVNGQNFENIQLSKAMEILR	(SEQ ID NO: 2)
dGRF4	EVNGQSLDHVTSKRALEILT	(SEQ ID NO: 16)
hPTP-BAS-1	SVNSVSLEGVSHHAAIEILQ	(SEQ ID NO: 17)
hPSD-95	SVNGVDLNRNASHEQAIALK	(SEQ ID NO: 18)
rLin-7-C	SVNGVSVEGEHHEKAVELLK	(SEQ ID NO: 19)
hDGL	EFNGINLRSATEQQARLIIG	(SEQ ID NO: 20)
	... * . . .	



Figure 9:

GRF4-RA domain

dgk-1a_ce_

GRF4

RaIGDS_h_

```

-----REDFE---IIRVFDGNNS---YRSQIS-----RNIVVAKHVSQVQRDAALR
--HHRILDFS---ATPDLPDQVLRVFKADQQS-----RYIMISKDTTAKVVITQAIR
SILVTSQDKAPSVISRVLKKNRDSAVASEYELVQLLPGERELTIPASANVFYAMDGASH
*
: : : : :
: : : : :

```

dgk-1a_ce_

GRF4

RaIGDS_h_

```

RFHI--NDTPERYIT-QVWGEVEEILED---PVP-----LRNVKRPPEGKRAQIFIR
EFAV--TATPDQYSLC-EVSVTPEGVIKQR-----RLP-----DQLSK--LADRIQLSGR
DFLLRHGEGPLLLHLASFVARLPQELLRVREEGAPFFGSRPQGGRLHGHCSSEEAPLAYR
* : * : * : * :

```

dgk-1a_ce_

GRF4

RaIGDS_h_

(SEQ ID NO: 26)

(SEQ ID NO: 2)

(SEQ ID NO: 27)

YYD-----

YYLKNNME--

SHGVHTRCG

:



6568 bp

/translation="MKPLAIPANHGVMGQEQKHSLPADFTKLHLTDSLHPQVTHVSSS
HSGCSITSDSGSSSLSDIYQATESEAGMDLGLPETAVDSEDDDDDEEDIERASDPLM
SRDIVRDCLEKDPIDRTDDIEQLLEFMHQLPAFANMTMSVRRELCAVMVFAVVERAG
TIVLNDGEELDSWSVILNGSVEVTYPDGKAEILCMGNSFGVSPTMDKEYMKGMVRTKV
DDCQFVCIAQQDYCRILNQVEKNMQKVEEAGEIVMVKEHRELDRTGTRKGHIVIKGTS
ERLTMHLVEEHSVVDPTFIEDFLLTYRTFLSSPMEVGKKLLEWFNDPSLRDKVTRVVL
LWVNNHFNDFEGDPAMTRFLEEFENNLEREKMGHRLRLNIACAAKAKRRLMTLTKPS
REAPLPFILLGGSEKGFIFVDSVDSGSKATEAGLKRGDQILEVNGQNFENIQLSKAM
EILRNNTLSITVKTNLVFKELLTRLSEEKRNAGPHLPKIGDIKKASRYSIPDLAVD
VEQVIGLEKVNKSKANTVGGRNKLKKILDKTRISILPQKPYNDIGIGQSQDDSIIVGL
RQTKHIPTALPVSGTLSSSNPDLLQSHHRLDFSATPDLPDQVLRVFKADQQSRYIMI
SKDTTAKENVVIAIREFAVTAATPDQYSLCEVSVTPEGVIKQRRLPDQLSKLADRIQLS
GRYYLKNMNETELCSDDEDAQELLRESQISLLQLSTVEVATQLSMRNFELEFRNIEPTE
YIDDLFKLRKSTSCANLKRFEVINQETFWVASEILRETNQLKRMKIIKHFIKIALHC
RECKNFNSMFAISGLNLAPVARLRTTWEKLPNKYEKLFQDLQDLFDPSRNMAYKRVN
LNSQNLQPPHPLFPVKKDLTFLHEGNDKVDGLVNFELRMIAKEIRHVGRMASVN
MDPALMFRTRKKKWRSLGSLSQGSTNATVLDVAQTGGHKKRVRRSSFLNAKKLYEDAQ
MARKVKQYLSNLELEMDDEESLQTLSQLCEPATNTLPKNPGDKKPVKSETSPVAPRAGS
QQAQSLPQPQQPPPAHKINQGLQVPAVSLYPSRKKVPVKDLPPFGINSQALKKIL
SLSEEGSLERHKKQAEDTISNASSQLSSPPTSPQSSPRKGYTLAPSGTVDNFSDSGHS
EISSRSSIVSNSSFDSPVSLHDERRQRHSVSIVETNLGMGRMERRTMIEPDQYSLGS
YAPMSEGRGLYATATVISSPSTEELSQQDQDRASLDAADSGRGSWTCSSGSHDNIQT
IQHQRSWETLPGHTHFDYSGDPAGLWASSSHMDQIMFSDHSTKYNRQNSRESLEQA
QSRASWASSTGYWGEDSEGDTGTIKRRGGKDVSI EAESSLTSVTTEETKPVMPMAHI
AVASSTTKGLIARKEGRYREPPPTPPGYIGIPITDFPEGHSHPARKPPDYNNVALQRSR
MVARSSDTAGPSSVQQPHGHPTSSRPVNKPQWHKPNESDPRAPYQSQGFSTEDEDE
QVSAV" | SEQ ID NO: 2]

63. .4562

BASE COUNT 1974 a 1400 c 1463 g 1731 t

ORIGIN

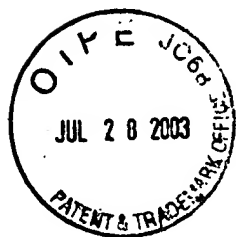
1 cttgccatcg tgagagattg gtacatgatg tgtaaatca gttcagcata tgtttctca
61 ttatgaaacc actagcaatc ccagctaacc atggagttat gggccagcag gagaaacact
121 cacttctcgc agatttcaca aaactgcac tcactgacag tctccacca caggtgaccc
181 acgtttcttc tagccattca ggatgtagta tcactagtga ttctgggagc agcagtctt
241 ctgatatcta ccagggccaca gaaagcgagg ctggtgatat ggacctgagt ggggtgccag
301 aaacagcagt ggattccgaa gacgacgacg atgaagaaga cattgagaga gcatcagatc
361 cctgatgag cagggacatt gtgagagact gcctagagaa ggacccaatt gaccggacag
421 atgatgacat tgaacaactc ttggaattta tgcaccaggt gcctgcttt gccaatatga
481 caatgtcagt gagggcagaa cctctgtctg tgatggtgtt cgcagtgtg gaaagagcag
541 ggaccatagt gttaatgat ggtgaagagc tggactcctg gtcagtgtt ctcaatggt
601 ctgtggaagt gacttatcca gatggaaaag cagaaatact gtgcatggga aatagtgtt
661 gtgtctctcc taccatggac aaagaataca tgaaggagat gatgagaaca aaggtggatg
721 actgccagtt tctctgcata gccagcaag attactgccg tatttcaat caagtagaaa
781 agaactgca aaaagttgaa gaggaaggag agattgtat ggtgaagaa caccgagaac
841 ttgatcgaac tggacaaga aaggacaca ttgtcatca ggttacctca gaaaggttaa
901 caatgcattt ggtggaagag cattcagtag tagatccaac attcatagaa gactttctgt
961 tgacctatag gactttctt tctagcccaa tggagtggtg caaaaagtt tggagtggt
1021 ttaatgacc gagcctcagg gataaggta cagggtagt attattgtg gtaataatc
1081 acttcaatga cttgaagga gatcctgcaa tgactcgatt ttagaagaa ttgaaaaca
1141 atctggaaag agagaaaatg ggtggacacc taaggctgtt gaatatcgcg tgtctgcta
1201 aagcaaaaag aagattgatg acgttaaca aaccatcccg agaagctcct ttgctttta
1261 tcttacttg aggtcttgag aagggtttg gaatcttgt tgacagtga gattcagga
1321 gcaagcaac tgaagcaggc ttgaaacggg gggatcagat attagaagta aatggccaaa
1381 accttgaaa cattcagctg tcaaaagcta tggaaattct tagaataac acacatttat

Figure 19a



1441 ctatcactgt gaaaaccaat ttattgtat taaagaact tctaacaaga ttgtcagaag
 1501 agaaaagaaa tgggtccccc caccctccta aaattggtga cattaaaaag gccagtgcgt
 1561 aciccattcc agatcttct gtatgttag aacaggtgat aggacttgaa aaagtgaaca
 1621 aaaaaagtaa agccaacact gtggaggaa ggaacaagct gaaaaagata ctgcacaaga
 1681 ctggatcag tatcttgcca cagaaccat acaatgatat tgggattgtt cagtctcaag
 1741 atgacagcat agtaggatta aggcagacaa agcacatccc aactgcattg cctgtcagt
 1801 gaaccttacc atccagtaat cctgatttat tgcagtcaca tcatcgatt ttagacttca
 1861 gtgctactcc tgaactgcca gatcaagtgc taagggtttt taaggctgat cagcaaagcc
 1921 gctacatcat gatcagtaag gacactacag caaagggaat ggctatcag gctatcagg
 1981 agttgtctgt tactgccacc ccggatcaat attcactatg tgaggctctt gtcacacctg
 2041 agggagtaat caaacaaga agacttcag atcagcttc caaactgca gacagaatac
 2101 aactgagtgg aaggtattat ctgaaaaaca acatggaaac agaaactctt tttcagatg
 2161 aagatgctca ggagttgtt agagagagtc aaatttccct ccttcagctc agcactgttg
 2221 aagttgcaac acagctctct atgcgaaatt ttgaactctt tcgcaacatt gaacctactg
 2281 aatatataga tgatttatt aaactcagat caaaaaccag ctgtgccaac ctgaagagat
 2341 ttgaagaagt cattaacag gaaacatttt gggtagcatt tgaattctc agaaaaaca
 2401 accagctgaa gaggatgaag atcattaagc atttcatcaa gatagcactg cactgtagg
 2461 aatgcaagaa tttaactca atgttgcaa tcatcagtg cctaaacctg gcaccagtgg
 2521 caagactcgc aacgacctgg gaaaaactt ccaataata cgaaaaaacta ttcaagatc
 2581 tccaagacct gttgatcct tccagaaaca tggcaaaata tctaatgtt ctcaatagtc
 2641 aaaaactaca accctccata atccctctat tccagttat caaaaaggat ctacacctcc
 2701 ttcacgaagg aaatgactca aaagttagac ggctgggtcaa tttgagaag ctaaggatga
 2761 ttgcaaaaaga aattcgtcac gttggccgaa tggcttcagt gaacatggac cctgccctca
 2821 tttcaggac tgggaagaag aaatggcgga gttggggtc tctcagccag ggtagtacaa
 2881 atgcaacagt gctagatgtt gctcagacag gtgtcataa aaagcgggta cgtcgtagt
 2941 cctttctcaa tgcacaaaag cttatgaag atgcccacat ggctcgaana gtgaagcagt
 3001 acccttccaa ttggagcta gaaatggac agagagagct tcaacatta tctctgcat
 3061 gtgagccagc aaccaacaca ttgcctaaga atcctgttga caaaaagcct gtcacatccg
 3121 agaccttcc agtagctcca agggcagggt cacaacagaa agctcagtc ctgccacag
 3181 ccagcagca gccaccacca gcacataaaa tcaaccaggg actacaggtt cccgcctgt
 3241 cctttatcc ttcacggaag aaagtgcctg taaaggatct cccacctttt ggcataaact
 3301 ctccacaagc tttaaaaaaa attcttctt tgtctgaaga aggaagtgtt gaacgtcaca
 3361 agaaacaggc tgaagataca atatcaaatg catcttcgca gctttctct cctctactt
 3421 ctccacagag ttctccaagg aaaggctata cttggctcc cagtgttact gtggataatt
 3481 ttccagattc tggtcacagt gaaattctt cagatccag tattgttagc aattcgtctt
 3541 ttgactcagt gccagtctca ctgcacgat agaggcgcca gaggcattct gtcagatcg
 3601 tggaaaacaa cctagggatg ggcaggatgg agaggcggac catgattgaa cctgatcat
 3661 atagcttggg gtctatgca ccaatgtcc agggccgagg ctatattgt acagctacag
 3721 taatttctc tccaagcaca gaggaactt cccaggatca gggggatcgc gcgtcactg
 3781 atgctgtga cagtggcctt gggagctgga cgtcatgctc aagtggctcc catgataata
 3841 tacagacgat ccagaccag agaagctggg agactcttcc attcgggcat actcactttg
 3901 attattcagg ggtatctgca gtttatggg catcaagcag ccatatggac caaattatg
 3961 ttctgatca tagcacaag tataacaggc aaatcaaaag tagagagagc ctgaacaag
 4021 cccagtcctg agcaagctgg gcgtcttcca caggttactg gggagaagac tcagaagggt
 4081 acacaggcac aataaagcgg aggggtggaa aggatgttc cattgaagcc gaaagcagta
 4141 gcctaacgct tgtgactac gaagaaacca agcctgtccc catgctgtcc cacatagctg
 4201 tggcatcaag tactacaaag gggctcattg cagaaagga gggcagggtat cgagagcccc
 4261 cggccacccc tcccggctac attggaattc ccttactga ctcttcagaa gggcactccc
 4321 atccagccag gaaaccgctg gactacaag tggcccttca gagatcgcg atggtcgac
 4381 gatcctccga cacagctggg ccttcatccg tacagcagcc acatgggcat cccaccagca
 4441 gcaggcctgt gaacaaacct cagtggcata aaccgaacga gtctgacctg cgcctcgccc
 4501 ctatcagtc ccaagggttt tccaccgagg aggatgaaga tgaacaagt tctgtgttt

Figure 19b



4561 gaggcacaga cttttctgga agcagagcga gccacctgaa aggagagcac aagaagacgt
 4621 cctgagcatt ggagccttg aactcacatt ctgaggacgg tggaccagtt tgcctccttc
 4681 cctgccttaa aagcagcatg gggtctcttc tccctctctt ccttccctt ttgcatgtga
 4741 aatactgtga agaaattgcc ctggcacttt tcagactttg ttgcttgaaa tgcacagtgc
 4801 agcaatcttc gagctccac tgttctgcc tggcacatca cacagtatca ttccaaattc
 4861 caagatcatc acaacaagat gattcactct ggctgcactt ctcaatgcct ggaaggattt
 4921 ttttaactct tcttttaga tttaaatcca gtcttagcac ttgatctcat tgggataatg
 4981 agaaaagcta gccattgaac tacttggggc cttaaccca ccaaggaaga caaagaaaaa
 5041 caatgaaatc ctttagtac agtgcctgtc cactgttta caatgtcctc ctttaaaaaa
 5101 aaaaaatgag tttaaagatt ttgtcagag agtaaatata tatccattta atgattacag
 5161 tattattta aaccttaagt aggggtgcc gcctgggttc tgaanaacca aatatgccgg
 5221 acagggtgtg gccacacca gaagacggga agacctggct tgtaccctg gcttcccatg
 5281 tcttctggt ctaccccgag aagtgcctta tcttggaagt atgaaatgt agccaattaa
 5341 taccaagaca cctcatctgc tcttcccca gtggatgggg tcttctgtta aaactgtttg
 5401 cacatggcca ggggagggaa ctaggacctt tgtctctgt ctgagcctta tggaggcagg
 5461 acggtgtcat tggcggatgt gtctgtctcc atgagatgg atggcaaacc ccatitttaa
 5521 gttatattc ttgatttt gtttaattag aggtgtaggt ttgttttt gtttttgtt
 5581 ttttttaag agaaacattt ataactggt agcattgcag tgaagcagc ttgggatgtt
 5641 ggagctaatt ccagctgttt atactgtct tcaagacag cctcccttta ttgaattggc
 5701 attagggaat aaacaagcct ttaacgtga taaaagatca aaacctggt tagacatgcc
 5761 agcctttgca aggcaggta gtaccaaag actaacctcc aagtggcttt atggacgctg
 5821 catatagaga aggcctaat gttagcaacca tctgtcaca gctgtattta accctataat
 5881 gactgaaatg accctccac tctattttg tgtgtttg cacagactcc ggaaaagtga
 5941 aggtgccaat tctgagtagt actcaaatgt gaggaactgc tggcttggga tttttttc
 6001 attaaattca gctgatcata ttgatcagta gataaacgta aatagcttca aatttaaaa
 6061 gtggaattgc agtgttttt cactgtatca aacaatgca gtgctttatt taataattct
 6121 ctctgtatc atggcatttg tctactgtct tattacattg tcaattatgc atttgaatt
 6181 ttacatgtaa tatgcattat ttgccagttt tattatag gctatggacc tcatgtgcat
 6241 atagaaagac agaaatctag cctaccaca agtgcacaa atgttatcta agcattaagt
 6301 aatttagaa cataggactg ctatctcag ttgctctgt gatgtcaagt gcagaatgta
 6361 caattaaactg gtgatttctt catactttg atactactg tacctgtatg tcttttagaa
 6421 agacattggt ggagtctgta tccctttgt attttaata caataattgt acatatggt
 6481 tatattttg ttgaagatgg tagaaatgta ctatgttat gcttctacat ccagtttgta
 6541 caagctggaa aataaataaa tataacat [SEQ ID NO: 1]

Figure 19c



801 b.p.

1/1
 ACT AAA GGG AAC AAA AGC TGG AGC TCC ACC GCG GTG GCG GCC GCT CTA GAA CTA GTG GAT
 T K G N K S W S S T A V A A A L E L V D
 L K G T K A G A P P R W R P L * N * W I
 * R E Q K L E L H R G G G R S R T S G S

31/11
 61/21
 CCC CCG GGC TGC AGG AAT TGA AGC GGT GGG AAG GAT GTC TCC GCT GAG GCA GAG AGC AGC
 P P G C R N S S G G K D V S A E A E S S
 P R A A G I Q A V G R M S P L R Q R A A
 P G L Q E F K R W E G C L R * G R E Q Q

91/31
 121/41
 AGC ATG GTG CCC GTG ACT ACA GAG GAA GCC AAA CCT GTC CCT ATG CCT GCC CAC ATA GCT
 S M V P V T T E E A K P V P M P A H I A
 A W C P * L Q R K P N L S L C L P T * L
 H G A R D Y R G S Q T C P Y A C P H S C

151/51
 181/61
 GTG AGC CCG AGC ACT ACC AAG GGA CTC ATC GCA CGG AAG GAA GGC AGG TAC CGG GAG CCG
 V T P S T T K G L I A R K E G R Y R E P
 * R R A L P R D S S H G R K A G T G S R
 D A E H Y Q G T H R T E G R Q V P G A A

211/71
 241/81
 CCT CCC ACA CCT CCA GGC TAC GTG GGC ATC CCC ATT GCC GAT TTC CCA GAA GGG CCT TGC
 P P T P P G Y V G I P I A D F P E G P C
 L P H L Q A T W A S P L P I S Q K G L A
 S H T S R L R G H P H C R F P R R A L P

271/91
 301/101
 CAC CCG GCC AGG AAG CCC CCG GAT TAC AAC GTG GCC CTG CAG CGG TCC CGC ATG GTG GCA
 H P A R K P P D Y N V A L Q R S R M V A
 T R P G S P R I T T W P C S G P A W W H
 P G Q E A P G L Q R G P A A V P H G G T

331/111
 361/121
 CCG CCC ACT GAG GCC CCG GCA CCG GGC CAG ACG CCG CCT GCA GCC GCA GCC AGC CCG CCG
 R P T E A P A P G Q T P P A A A A S R P
 G P L R P R H R A R R R L Q P Q P A G R
 A H * G P G T G P D A A C S R S Q P A G

391/131
 421/141
 GGC AGC AAG CCA CAG TGG CAC AAG CCC AGC GAC GCA GAC CCA CGC CTC GCG CCC TTC CAG
 G S K P Q W H K P S D A D P R L A P F Q
 A A S H S G T S P A T Q T H A S R P S S
 Q Q A T V A Q A Q R R R P T P R A L P A

451/151
 481/161
 CCG CAG GCT TCG CAG GAG CCG AGG AGG ACG AAG ATG AAC AAG TGT CTG CTG TTT GAG GCG
 P Q A S Q E R R R T K M N K C L L F E A
 R R L R R S G G G R R * T S V C C L R R
 A G F A G A E E D E D E Q V S A V * G A

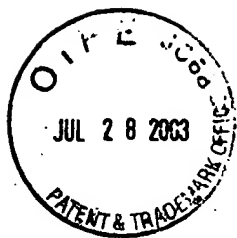
511/171
 541/181
 CAG GCT CCT TGA TCC ACA GTG AGC CAC CCA AAG GAG AGC ACA AGA AGA CGT CCC AAG CCT
 Q A P * S T V S H P K E S T R R R P K P
 R L L D P Q * A T Q R R A Q E D V P S L
 G S L I H S E P P K G E H K K T S Q A L

571/191
 601/201
 TGG AGC CTT GGC ACG CAC ATC TGA GGA TGG TGG ACC AGT TTG CCT CCT TCC CTG CCT TAA
 W S L G T H I * G W W T S L P P S L P *
 G A L A R T S E D G G P V C L L P C L K
 E P W H A H L R M V D Q F A S F P A L K

631/211
 661/221
 AGC AGC ATG GGG CTT CTT CTC CCC TTC TTT CCT TTT CAT GTG AAA TAC TGT GAA
 S S M G L L L P F F L S P L H V K Y C E
 A A W G F F S P S S F P L C M * N T V K
 Q H G A S S P L L P F P F A C E I L * R

691/231

Figure 19d



721/241
GAA ATT GCC CTG GCA CTT TGC AGA CTT GTT GCT TGA AAT GCA CAG CCC AGC AGC CCC TGA
B I A L A L C R L V A N A Q P S S P
K L P W H P A D L L L E M H S P A A P E
N C P G T L Q T C C L K C T A Q Q P L S
751/251
781/261
GCT GCT GCC TGC CAC GTC ACG [SEQ ID NO: 3]
A A A C H V T [SEQ ID NO: 4]
L L P A T S [SEQ ID NO: 5]
C C L P R H [SEQ ID NO: 6]

FIGURE 19E